Complete and assembled genome sequence of an NDM-5 and CTX-M-15 producing Escherichia coli sequence type 617 isolated from wastewater in Switzerland

Zurfluh, Katrin ; Stevens, Marc J A ; Stephan, Roger ; Nüesch-Inderbinen, Magdalena

Abstract: OBJECTIVES: Carbapenem-resistant Escherichia coli have emerged worldwide and represent a major challenge to effective healthcare management. Here we report the genome sequence of an NDM-5- and CTX-M-15-producing E. coli belonging to sequence type 617 isolated from wastewater treatment plant effluent in Switzerland. METHODS: Whole-genome sequencing of E. coli 657SK2 was performed using Pacific Biosciences (PacBio) single-molecule real-time (SMRT) technology RS2 reads (C4/P6 chemistry). De novo assembly was carried out using Canu 1.6, and sequences were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). RESULTS: The genome of E. coli 657SK2 consists of a 4.9-Mbp chromosome containing blaCTX-M-15, genes associated with virulence [fyuA, hlyE, the pyelonephritis-associated pili (pap) gene cluster and the yad gene cluster], the copper resistance gene pco, and genes associated with resistance to quaternary ammonium compound (QAC) disinfectants (emrA, mdfA and sugE). A 173.9-kb multidrug resistance IncFII-FIA-FIB plasmid was detected harbouring aadA2, aadA5, blaNDM-5, blaOXA-1, cat, drfA, drfA17, the mph(A)-mrx-mphR cluster, the tetA-tetC-tetR cluster, and the virulence genes iutA and ylpA. CONCLUSIONS: The genome sequence of E. coli 657SK2 provides information on resistance mechanisms and virulence characteristics of pathogenic E. coli harbouring blaNDM-5 and blaCTX-M-15 that are spreading into the environment via urban wastewater.

DOI: https://doi.org/10.1016/j.jgar.2018.08.015

Complete and assembled genome sequence of an NDM-5 and CTX-M-15 producing *Escherichia coli* sequence type 617 isolated from wastewater in Switzerland

Katrin Zurfluh[1], Marc J.A. Stevens[1], Roger Stephan[1*] Magdalena Nüesch-Inderbinen[1]

[1] Institute for Food Safety and Hygiene, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland

*Corresponding author:
Roger Stephan, Institute for Food Safety and Hygiene, Vetsuisse Faculty, University of Zurich, Winterthurerstrasse 272, CH-8057 Zurich, Switzerland.
Phone +41 44 635 86 51, Fax +41 44 635 89 08, e-mail stephanr@fsafety.uzh.ch

Abstract

Objectives: Carbapenem-resistant *Escherichia coli* have emerged worldwide and represent a major challenge to effective health care management. Here, we report the genome sequence of an NDM-5 and CTX-M-15 producing *E. coli* belonging to sequence type 617, isolated from wastewater treatment plant effluent in Switzerland.

Methods: Whole genome sequencing of *E. coli* 657SK2 was performed using Pacific Biosciences (PacBio) single-molecule real-time (SMRT) technology RS2 reads (C4/P6 chemistry). De novo assembly was carried out using CANU 1.6, and sequences were annotated using the NCBI Prokaryotic Genome Annotation Pipeline.

Results: The genome of *E. coli* 657SK2 consists of a 4.9-Mbp chromosome containing $bla_{\text{CTX-M-15}}$, genes associated with virulence (*fyuA*, *hlyE*, the pyelonephritis-associated pili *(pap)* gene cluster, and the *yad* gene cluster), the copper resistance gene *pco*, and genes associated with resistance to quaternary ammonium compound (QAC) disinfectants, *emrA*, *mdfA*, and *sugE*. A 173,9-kb multidrug resistance IncFII-FIA-FIB plasmid was detected harbouring *aadA2*, *aadA5*, $bla_{\text{NDM-5}}$, $bla_{\text{OXA-1}}$, *cat*, *drfA*, *drfA17*, the *mph(A)-mrx-mphR* cluster, the *tetA-tetC-tetR* cluster, and the virulence genes *iutA,* and *ylpA*.

Conclusions: The genome sequence of *E. coli* 657SK2 provides information on resistance mechanisms and virulence characteristics of pathogenic *E. coli* harbouring $bla_{\text{NDM-5}}$ and $bla_{\text{CTX-M-15}}$ that are spreading into the environment via urban wastewater.

Keywords

E. coli ST617, $bla_{\text{NDM-5}}$, $bla_{\text{CTX-M-15}}$, genome analysis

The dissemination of carbapenemase-producing Enterobacteriaceae (CRE) constitutes a challenge to effective healthcare management worldwide. New Delhi metallo-β-lactamases (NDM) positive *E. coli* have been identified in infected or colonised patients in the clinical setting and sporadically from animal or environmental sources [1]. The NDM-5 variant, which is infrequently detected worldwide, differs from NDM-1 by enhanced hydrolytic activity against carbapenems and the point mutations at positions 88 (Val → Leu) and position 154 (Met → Leu).

E. coli 657SK2 was isolated from wastewater treatment plant effluent which is released into the river Rhine near Basel, Switzerland in December 2015 [2]. DNA extraction was performed with the Wizard® Genomic DNAPurification Kit according to the manufacturers protocol (Promega AG, Dübendorf, Switzerland). The genome was sequenced at the Functional Genomics Center Zurich (FGCZ), Switzerland, using two single-molecule real-time (SMRT) cells on a PacBio RS II (Pacific Biosciences, Menlo Park, CA, USA). The raw reads were filtered using the RS Filter Only protocol in the SMRT-portal (Pacific Biosciences) using standard settings. A total of 84,248 reads with an average length of 10,450 bp were selected, corresponding to 880,442,668 sequenced basepairs and a genome coverage of approximately 175 fold. The reads were assembled using Canu 1.6 [3] with the option "-pacbio-filtered" and an estimated genome size of 5.0 Mbp. The Canu output consisted of 4 contigs which were further polished in CLC workbench 7 (CLC, Aarhus, Denmark), eventually resulting in one chromosome, one phage and one plasmid encoding contig. The chromosomal origin of replication was identified using DoriC 5.0 and plasmidal origin of replication were determined by PlasmidFinder 1.3. The start of the chromosome was set 9 bp upstream of the first DnaA box in the origin of replication region. The genome was annotated by the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (GPAP) server.

68 The chromosome of *E. coli* ST617 657SK2 has a G+C content of 49.9 % and consists of a
69 4,889,056-bp chromosome. One extra-chromosomal element of 173,883 bp was identified
70 and designated p657SK2. The Prokaryote Genome Annotation Pipeline (PGAP) at NCBI
71 predicted 5250 genes and 5125 protein-coding sequences. A total of 296 frame-shifted genes
72 were identified by PGAP and manual analysis revealed that 128 frame shifts were caused by
73 homopolymer sequencing errors. These 128 frameshifts were corrected using the genome of
74 *E. coli* K12 substrain MG1655 (accession number CP027060.1) as leading sequence.
75 The chromosome of *E. coli* 657SK2 contains the ß-lactam resistance gene $bla_{\text{CTX-M-15}}$,
76 virulence associated genes *fyuA*, *hlyE*, the pyelonephritis-associated pili *(pap)* gene cluster,
77 and the *yad* gene cluster, as well as the copper resistance gene *pco*, and genes associated with
78 resistance to quaternary ammonium compound (QAC) disinfectants, *emrA*, *mdfA*, and *sugE*.
79 Plasmid p657SK2 is a multireplicon IncFII-FIA-FIB plasmid assigned to the FAB formula
80 F31:A4:B1 by the IncF replicon typing scheme (http://cge.cbs.dtu.dk/services/pMLST/) [4].
81 In addition, the presence of a second FII allele (F36) was observed, which is composed of
82 *repA2*, *repA6* and truncated *repA1*. Plasmid p657SK is a multidrug resistance (MDR)
83 plasmid encoding the following resistance genes: aminoglycosides (*aadA2* and *aacA5*), ß-
84 lactams ($bla_{\text{NDM-5}}$, $bla_{\text{OXA-1}}$), bleomycin (*ble*), chloramphenicol (*cat*), macrolides (*mph(A)*-
85 *mrx-mphR* cluster), tetracycline (*tetA-tetC-tetR* cluster), trimethoprim (*drfA* and *drfA17*). The
86 plasmid further contains virulence genes *iutA, ylpA*, and two truncated copies of the QAC
87 resistance gene *emrE*.
88 Genes for four toxin/antitoxin systems were detected (*ccdA*/*ccdB*, *hok*, *pemI*/*pemK*, and
89 *vapB*/*vapC*).
90 Plasmid p657SK2 closely resembles the $bla_{\text{CTX-M-15}}$ encoding plasmids pCA14 (CP009231)
91 and pCA28 (CP009232) from *E. coli* causing community-acquired infections in the USA [5],
92 each with 85% coverage and 99% identity. The $bla_{\text{NDM-5}}$ gene is located downstream of

93 truncated IS*Aba125* and upstream of *ble*. The genetic environment of this highly conserved
94 arrangement is on a 22.7 kbp region which is bracketed by two IS26 elements and constitutes
95 a putative composite transposon containing a class I integron, aminoglycoside resistance
96 genes *aadA2* and *aacA5, drfA, drfA17*, and *sul1,* and an insertion sequence common region 1
97 (ISCR1) replicase (Figure 1). The *intI1* gene downstream of $bla_{\text{NDM-5}}$ is truncated by IS26
98 (Figure 1). The upstream region of the $bla_{\text{NDM-5}}$ gene contains a resistance gene locus
99 consisting of IS6100 and the *mph(A)-mrx-mphR* cluster.
100 The occurrence of $bla_{\text{NDM-5}}$ in this genetic constellation suggests a high potential of this
101 carbapenemase gene to disseminate to other plasmids and species.
102 In summary, we report the draft genome of a carbapenem resistant *E. coli* ST617 isolated
103 from wastewater treatment plant effluent. The occurrence of NDM-5 producing *E. coli* in
104 treated urban wastewater that is released into one of the European continent's major rivers is
105 cause for concern. Environmental pollution with multidrug resistant, NDM-producing *E. coli*
106 harbouring genes associated with virulence and resistance to heavy metals and disinfectants
107 poses a threat to human and animal health.
108
109 Sequence and annotation data of the genome have been deposited at GenBank under
110 accession numbers **CP027701** (chromosome), and **CP027703** (p657SK2). This is the first
111 version of this genome.
112
113 **Funding**
114 This work was supported by funding from the University of Zurich.
115
116 **Competing interests**
117 None declared.

Ethical approval

Not required.

References

[1] Johnson AP, Woodford N. Global spread of antibiotic resistance: the example of New Delhi metallo-β-lactamase (NDM)-mediated carbapenem resistance. J Med Microbiol 2013; 62:499-513.

[2] Zurfluh K, Bagutti C, Brodmann P, Alt M, Schulze J, Fanning S et al. Wastewater is a reservoir for clinically relevant carbapenemase- and 16s rRNA methylase-producing Enterobacteriaceae. Int J Antimicrob Agents 2017; 50:436-440.

[3] Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 2017; 27:722-736.

[4] Carattoli A, Zankari E, Garcìa-Fernandez A et al. PlasmidFinder and pMLST: in silico detection and typing of plasmids. Antimicrob Agents Chemother 2014; 58:3895-3903.

[5] Li JJ, Spychala CN, Hu F, Sheng JF, Doi Y. Complete nucleotide sequences of *bla*$_{\text{CTX-M}}$-harboring IncF plasmids from community-associated *Escherichia coli* strains in the United States. Antimicrob Agents Chemother 2015; 59:3002-3007.

Figure legend

Linear map of the 22.7 kbp region of p657SK2 containing *bla*$_{NDM-5}$. Antimicrobial resistance genes are coloured in red, insertion sequences (IS) are shown in dark or light blue, resistance to quaternary ammonium compound (QAC) disinfectants are shown in green, other genes are shown in burgundy.

aacA5, aminoglycoside N(6')-acetyltransferase gene*;* *bla*$_{NDM-5}$, gene coding for the New Delhi metallo-ß-lactamase; *ble*, bleomycin resistance gene; *dfrA*, dihydrofolate reductase gene; *dosP*, gene for oxygen sensor protein; *dsbC*; disulfide bond isomerase II gene; Δ*emrE*, truncated quaternary ammonium compound-resistance gene; *intI1*, integron integrase gene; ISCR-1, gene for the rolling circle replicase of the insertion sequence common region 1 transposase; *srpC*, gene for chromate transport protein; *sul1*, sulfonamide resistance gene; *trpF*, phosphoribosylanthranilate isomerase gene.